Serial Attached SCSI storage technology

2nd Edition

Abstract 2

Transition from parallel to serial SCSI protocol 2

SAS terminology 2

SAS technology 3
- SAS devices 3
  - Initiators 4
  - Expanders 4
  - Targets 4
- Differential signaling 5

SAS protocol evolution 6
- SAS-1 6
- SAS-2 6
- SAS-2.1 6
  - Active cables 6
  - Storage power management 7

SAS/SATA interoperability 8

Cabling and connectors 9
- Mini SAS 4x cable connectors and receptacles 9
- Mini SAS 8x cable connectors 11

SAS topologies 11
  - Internal 11
  - External 13
  - Multi-node clusters 14

Zoning 14

SAS performance 16

For more information 18

Call to action 18

Downloaded from www.Manualslib.com manuals search engine

Abstract

This technalagy brief describes Serial Attached SCSI-2 (SAS-2) technalagy, the evalutian af the SAS standard, interaperability with Serial Advanced Technalagy Attachment (SATA) devices, enhancements ta SAS-2 cabling and cannectars, and SAS tapalagies and zaning.

Transition from parallel to serial SCSI protocol

The ariginal SCSI standard was develaped in 1981 ta pravide a camman interface that cauld be used acrass all peripheral platfarms and system applicatians, such as Redundant Array af Independent Disks (RAID) starage. Since that time, there have been seven generatians af the parallel SCSI pratacal. Each generatian daubled the bandwidth af the previaus ane, primarily by daubling the bus clack frequency. But as the bus frequency was increased with each new generatian, sa did the negative impact af bus cantentian, signal degradatian, and signal skew—slight signal delays fram ane wire trace ta the next. After the develapment af Ultra320 SCSI with a bandwidth af 320 MB/s per channel, further bandwidth impravements ta parallel SCSI cauld nat accur withaut develaping new and expensive technalagies.

In 2001, HP (Campaq), IBM, LSI Lagic, Maxtar, and Seagate faunded the Serial Attached SCSI Warking Graup ta define the rules far exchanging infarmatian between SCSI devices using a serial intercannect (SAS). SAS was later transferred ta the InterNatianal Cammittee far Infarmatian Technalagy Standards (INCITS) T10 ta became an American (ANSI) and internatianal (ISO/IEC) standard.

SAS inherits its cammand set fram parallel SCSI, frame farmats and full duplex cammunicatian fram Fibre Channel, and it uses the SATA interface far campatibility and investment pratectian. The SAS architecture salves the parallel SCSI prablems af bus cantentian, clack skew, and signal degradatian at higher signaling rates, thereby praviding perfarmance headraam ta meet enterprise starage needs far years ta came.

SAS terminology

Table 1 pravides a list af terms related ta SAS technalagy ta aid in understanding the cancepts described in this paper.

Table 1. SAS terminology

Term	Definitian
ANSI	American Natianal Standards Institute
Damain	An I/O system cansisting af a set af SAS devices that cammunicate with ane anather by means af a service delivery subsystem
Enterprise-class devices	SAS drives that pravide maximum reliability, highest perfarmance, scalability, and errar management far use with uncanstrained I/O warklaads in missian-critical applicatians
Entry level drives	SATA drives with the lawest unit cast that pravide a basic level af reliability and perfarmance far nan-missian-critical environments
Expander	A device that functians as a switch to attach ane ar mare initiators to ane ar mare targets
Initiator	A device cantaining SSP, STP, and/ar SMP initiator parts in a SAS damain

Downloaded from www.Manualslib.com manuals search engine

ISO/IEC	International Organization for Standardization/International Electrotechnical Commission
Midline devices	SAS and SATA drives that provide larger capacity, greater reliability, improved resistance to rotational and operational vibration than Entry level drives, making them better suited for use in multi-drive configurations.
Phy	The mechanism that contains a transceiver which electrically interfaces to a physical link. Phy is a common abbreviation for the physical layer of the OSI model.
Physical link	Two differential signal pairs, one pair in each direction, that connect two physical phys
SAS address	The identifier of an initiator port, a target port, or an expander device
Serial ATA Tunneling Protocol (STP)	A protocol used to communicate with SATA drives
Serial Management Protocol (SMP)	A protocol used to communicate with SAS expanders
Serial SCSI Protocol (SSP)	A protocol used to communicate with SAS drives
Service delivery subsystem	The part of a SAS I/O system that transmits information between a SAS initiator port and a SAS target port
Subtractive routing	A routing technique used when a device is not able to find other devices in the same sub-branch. This will pass the request to a different branch altogether.
Table routing	Table routing is used for identifying devices connected to the expanders connected to a device's own phy.
Target	An end device such as a SAS hard disk drive, SATA hard disk drive, or SAS tape drive
Training	The process of adapting equalization circuitry in a receiver device to an incoming transmission pattern
Virtual phy	Contains a vendor-specific interface to another virtual phy
Wide link	A group of physical links that attaches a wide port to another wide port
Wide port	A port that contains more than one phy
Zone group	A set of phys in a zone that all have the same access permission
Zoned portion of a service delivery subsystem (ZPSDS)	A group of zoning expander devices that cooperate to control access between phys

SAS technology

SAS is a point-to-point architecture that transfers data to and from SCSI storage devices by using serial communication (one bit at a time). SAS devices and the differential signaling method they use to achieve reliable, high-speed serial communication are described in this section.

SAS devices

There are three types of SAS devices: initiators, targets, and expanders. An initiator device is a host bus adaptor (HBA), or controller. The initiator is attached to one or more targets—SAS hard disk drives, SATA hard disk drives, and SAS tape drives—to form a SAS domain. Expanders are low-cost, high-speed switches that scale the number of targets attached to an initiator, thereby creating a larger

Downloaded from www.Manualslib.com manuals search engine

SAS domain (Figure 1). Each SAS device has a unique worldwide name (SAS address) assigned at manufacturing to simplify its identification in a domain

Initiators

SAS initiators have multiple ports for connection to internal and/or external targets. Each initiator port can have a single physical link (a narrow port) or 2, 4, or 8 physical links (a wide port). SAS initiator ports can be connected to separate domains for fail-over redundancy.

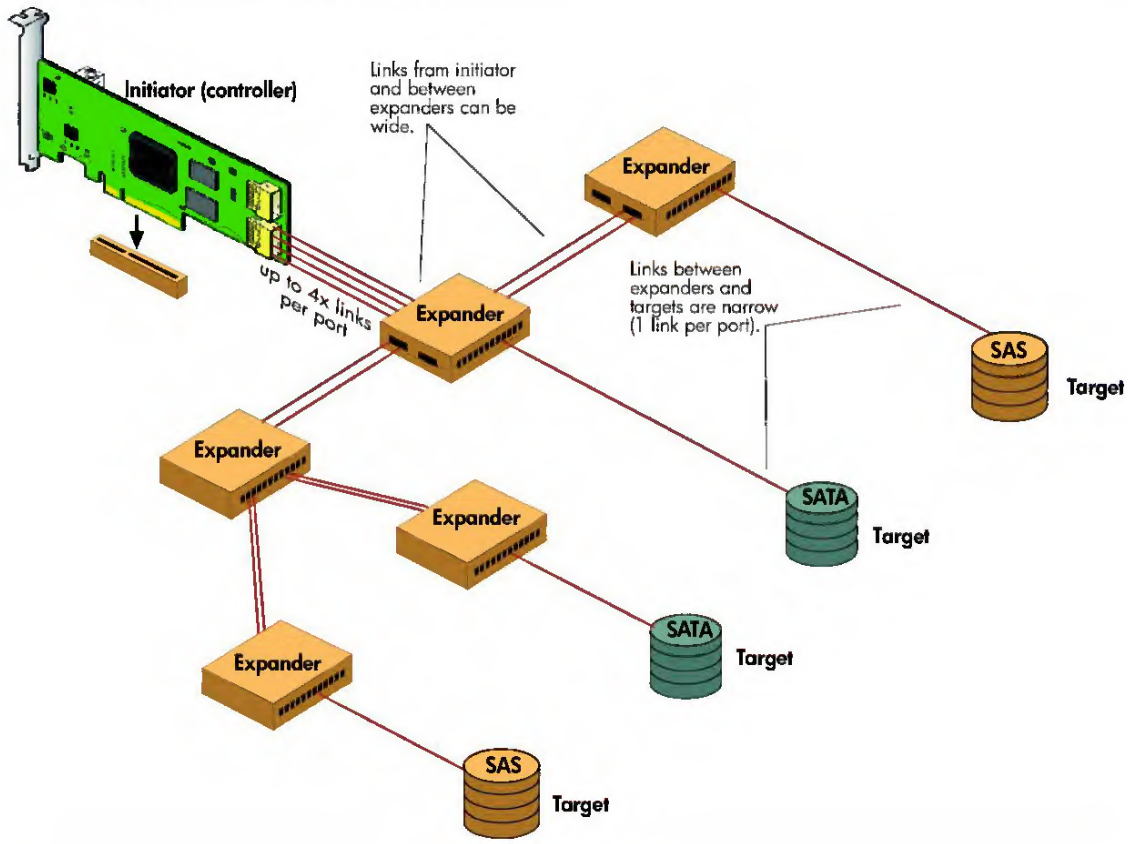
Expanders

Expanders establish connections between initiators, targets, and other expanders by receiving commands and data in one port and routing them to another port based on the SAS address of the target. Expanders use three routing methods—direct, table, and subtractive. An expander uses direct routing to forward commands and data to targets directly attached to the expander. An expander uses table routing to forward commands and data to another expander. When an expander receives an address that it does not recognize, it uses subtractive routing to forward the commands and data to another expander that does recognize the address.

Targets

SAS hard drives (enterprise-class and midline devices) have two narrow ports. Each port can reside in a different SAS domain to provide fail-over redundancy and load balancing. SAS hard drives leverage a common electrical and physical connection interface with SATA hard drives. However, SATA hard drives, including solid state drives, have a single narrow port.

Figure 1. Example of SAS devices in a domain. The number of initiators and targets allowed in a domain is limited only by the size of the expanders' routing tables.



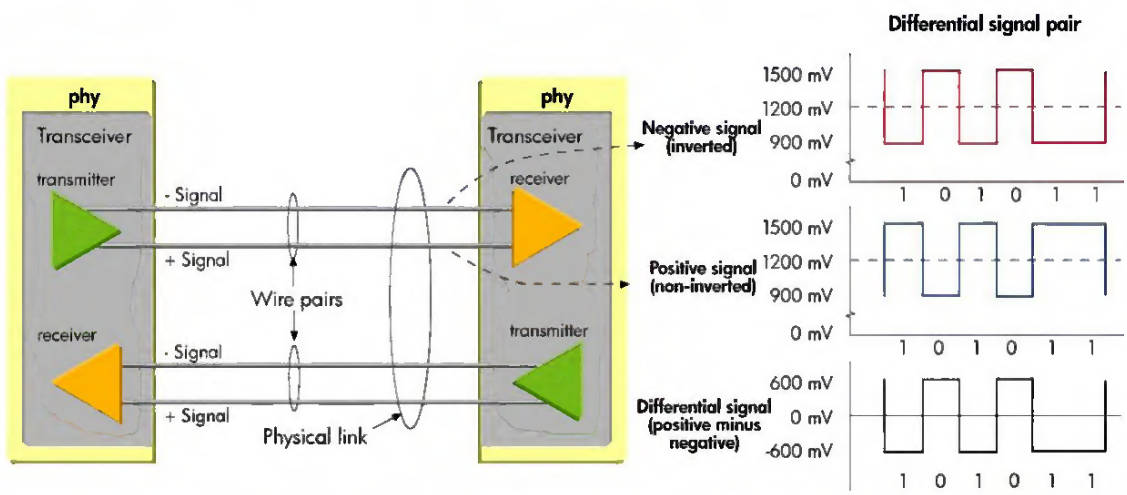
Downloaded from www.Manualslib.com manuals search engine

Differential signaling

All SAS devices have connection points called ports. One or more transceiver mechanisms, called phys, are located in the port of each SAS device. A physical link, consisting of two wire pairs, connects the transmitter of each phy in one device's port to the receiver of a phy in another device's port. The SAS interface allows the combination of multiple physical links to create two (2x), 3x, 4x, or 8x connections per port for scalable bandwidth. A port that has one phy is described as "narrow" while a port with two to four phys is described as "wide."

SAS uses differential signaling to transfer data over a physical link (Figure 2), which reduces the effects of capacitance, inductance, and noise experienced by parallel SCSI at higher speeds. SAS communication is full duplex, which means that each phy can send and receive information simultaneously over the two wire pairs.

Figure 2. In differential signaling, positive minus negative equals 1500–900 = 600mV or 900–1500 = -600mV.



The physical link rates for SAS and SATA technologies are listed in Table 2.

Table 2. Physical link rates per direction

Physical link rate	Generation	Bandwidth	4x bandwidth
1.5 Gbps	SAS-1, SAS-1.1, SATA Revision 1.0	150 MB/s	600 MB/s
3 Gbps	SAS-1, SAS-1.1, SATA Revision 2.0	300 MB/s	1200 MB/s
6 Gbps	SAS-2, SAS-2.1, SATA Revision 3.0	600 MB/s	2400 MB/s

Downloaded from www.Manualslib.com manuals search engine

SAS protocol evolution

SAS-1

The speed of the first-generation SAS (SAS-1) link is 3.0 gigabits per second (Gb/s). The original SAS standard defined two classes of expanders: edge expanders and fanout expanders. An edge expander attaches directly to targets, or to another edge expander, which reduces its complexity (and cost) by constraining the size of its routing table. Edge expanders can also be connected together to form an edge expander device set, allowing the set to address up to 128 devices. A fanout expander allows multiple edge expanders or edge expander sets to communicate with each other; therefore, it has to maintain a more extensive routing table.

SAS-2

The second-generation SAS (SAS-2) link speed doubles the physical link rate to 6.0 Gb/s. SAS-2 eliminates the distinction between fanout and edge expanders by replacing them with self-configuring expanders. SAS-2 also enables zoning for enhanced resource deployment flexibility, security, and data traffic management. SAS-2 is backward compatible with SAS-1.

Each SAS device (initiator, target, or expander) may support one or more SAS communication speeds. If any two linked devices support multiple speeds, the highest speed will be used if performance is critical or a slower speed may be used if reliability is more important. During this speed negotiation process, linked SAS devices determine their mutually supported SAS speed(s), and other speed-related options, such as transmission amplitude, slew rate, de-emphasis, and spread spectrum clocking.

After the speed negotiation process is complete, the SAS-2 standard provides a time interval, or training window, during which the linked devices exchange predefined signals to test the link using the selected speed(s) and speed-related options.

SAS-2.1

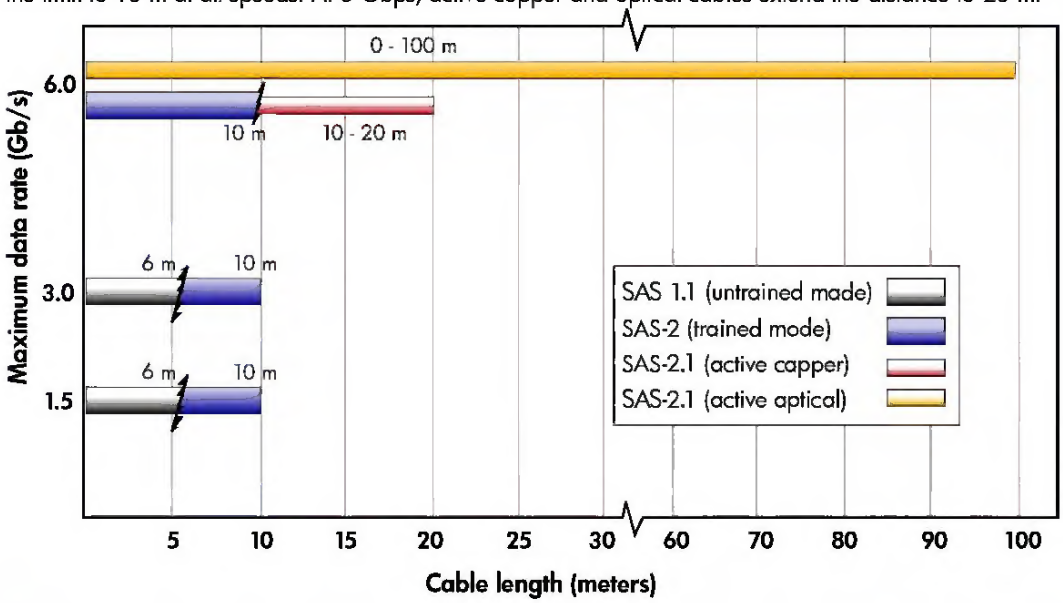
The enhancement to the SAS-2 standard, SAS-2.1, will define active cables, storage power management, and additional connectors (see "Cabling and connectors"). Also, SAS-2.1 splits out the protocol layer into a separate standard, SAS Protocol Layer (SPL).

Active cables

To help reduce cable weight, improve cable management, and improve airflow, the SAS 2.1 standard implements thinner cables with active circuitry. Active circuitry includes built-in drivers and repeaters, along with an equalizer. The equalizer removes inter-symbol interference (ISI), a form of signal distortion. The drivers and repeaters reduce the signal-to-noise ratio (SNR) by boosting the received signal and by reducing near-end crosstalk (NEXT). NEXT occurs when two wires are close enough for the signal traveling in one wire to interfere with the signal traveling in the other. Cable designers reduced this phenomenon by placing a low-power equalizing filter inside the cable to compensate for the dielectric and conductor losses. Economical, low-power, low-latency, active cables enable high data transfer rates using thinner wire gauges over longer cable runs (Figure 3).

Downloaded from www.Manualslib.com manuals search engine

Figure 3. SAS 1.1 at 1.5 Gbps and 3.0 Gbps in untrained mode has a 6-m limit. SAS-2 in trained mode raises the limit to 10 m at all speeds. At 6 Gbps, active copper and optical cables extend the distance to 20 m.



Storage power management

SAS-2.1 devices will be able to turn off SAS physical links when they are idle. Each SAS transceiver consumes about 200 mW. Therefore, the total power savings from a dual-ported drive with two transceivers and a controller (or attached SAS expander) with two transceivers is a little less than 1 W. Table 3 includes an example of power savings for a 9-W small form factor (SFF) SAS drive.

Future SAS and SATA disk drives will include more standardized power management features. Today's drives support the Active and Stopped states. The Stopped state is rarely used due to its long recovery time.

Table 3. Power management state example for a 9-W SFF SAS drive

State	Description	Commands processed	Power savings	Recovery time
Active	Fully active	Yes	None	None
Idle	Stop clocks to idle circuitry	Yes	~1 W	< 1 ms
Park heads	Retract the heads from the media, reducing drag	Automatic state change	~2 W	1 sec
Low-rpm standby	Retract heads, spin down to a lower rpm	Automatic state change	~4 W	7 sec
Standby	Retract heads, spin down to 0 rpm, cache still powered	Automatic state change	~7 W	15-20 sec
Stopped	Same as Standby	No	~7 W	15-20 sec
Sleep	Everything off except wakeup circuitry	No	~9 W	30 sec

Downloaded from www.Manualslib.com manuals search engine

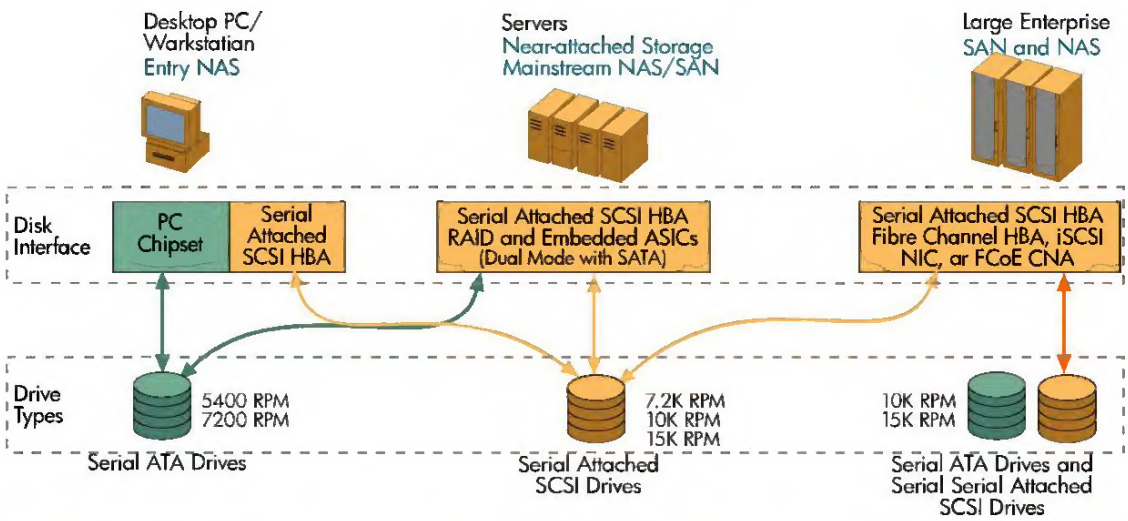
SAS/SATA interoperability

Using the SAS interfoce, SAS drives, SATA drives, or o mix of both SAS ond SATA drives will function in the some storoge enclosure. With this brood ronge of storoge solutions (Figure 4), IT monogers con choose storoge devices bosed on reliobility, performonce, ond cost. Becouse the SAS orchitecture feotures o proven SCSI commond set, odvonced commond queuing, ond odvonced verificotion/error correction, SAS is the ideol solution for mission-criticol enterprise storoge opplicotions.

SAS supports three protocols to hondle communicotions with vorious devices:

- Seriol Monogement Protocol (SMP) – monoges the point-to-point topology of exponders ond enclosures
- Seriol SCSI Protocol (SSP) – focilitotes communicotion with SAS devices ond existing SCSI softwore
- SATA Tunneling Protocol (STP) – ollows SAS controllers to communicote with SATA devices through exponders

Figure 4. The SAS orchitecture provides flexible solutions ronging from entry level NAS in desktop environments to mission-criticol RAID systems thot complement Fibre Chonnel technology.



Downloaded from www.Manualslib.com manuals search engine

Cabling and connectors

Mini SAS 4x cable connectors and receptacles

Mini SAS 4x connectors and receptacles (Figure 5) are replacing SAS 4x connectors and receptacles in new designs. Mini SAS connectors contain ground pins that can be used for power in active cables.

Figure 5. The internal Mini SAS 4i connector (left) and receptacle (right) are replacing SAS 4x connectors and receptacles.

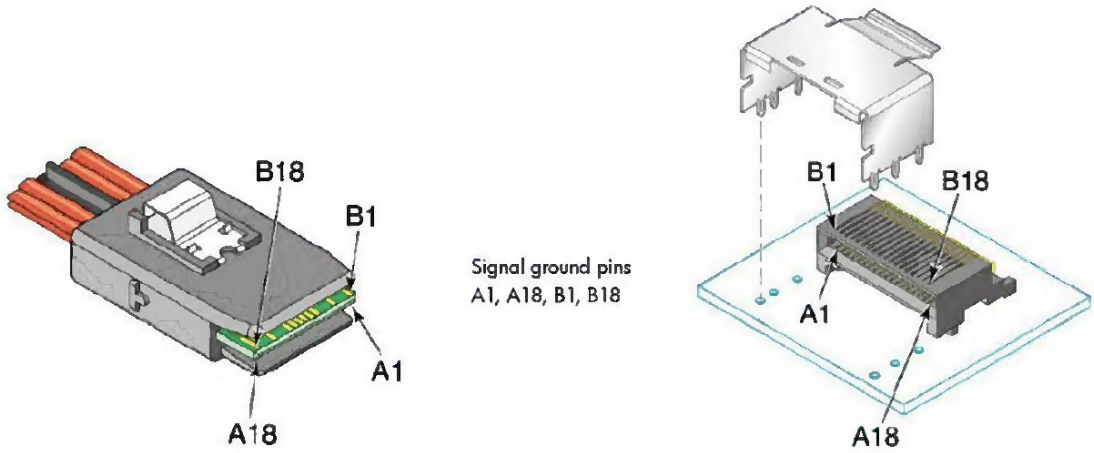
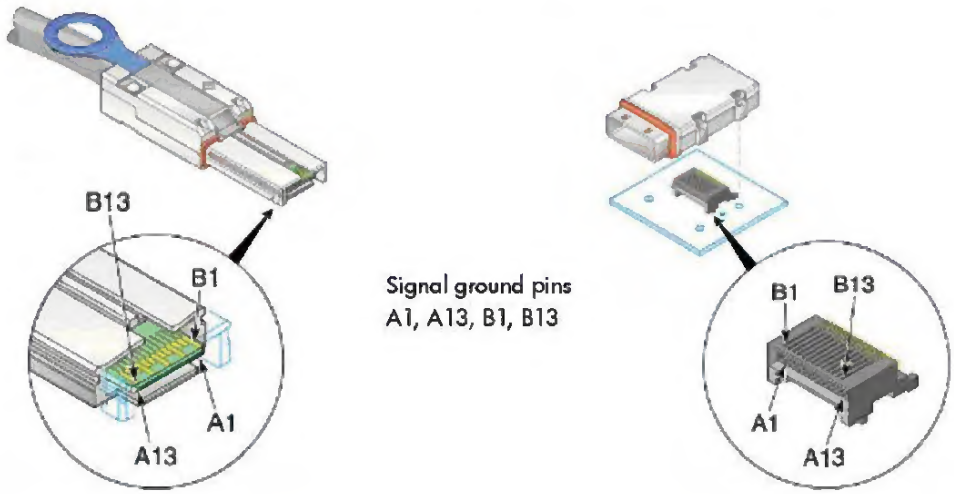


Figure 6 shows the Mini SAS 4x (external) cable plug connector and receptacle. Mini SAS 4x cable connectors attach to end devices and enclosure universal ports, but they are specific to enclosure out and in ports. The SAS-2 receptacles are universal, thus they do not have table routing or subtractive routing restrictions. The receptacles also have a reverse key (not illustrated) to accept cables longer than 6 m as well as shorter SAS-1 cables 6 m or less.

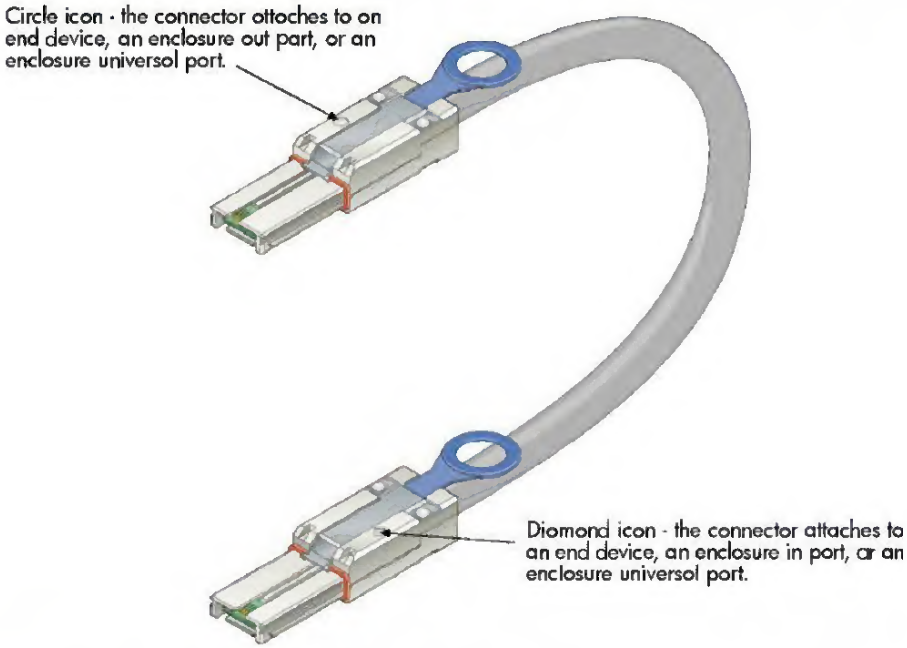
Figure 6. The external Mini SAS 4x connector is specific to enclosure in and out ports, but the receptacle is universal.



Downloaded from www.Manualslib.com manuals search engine

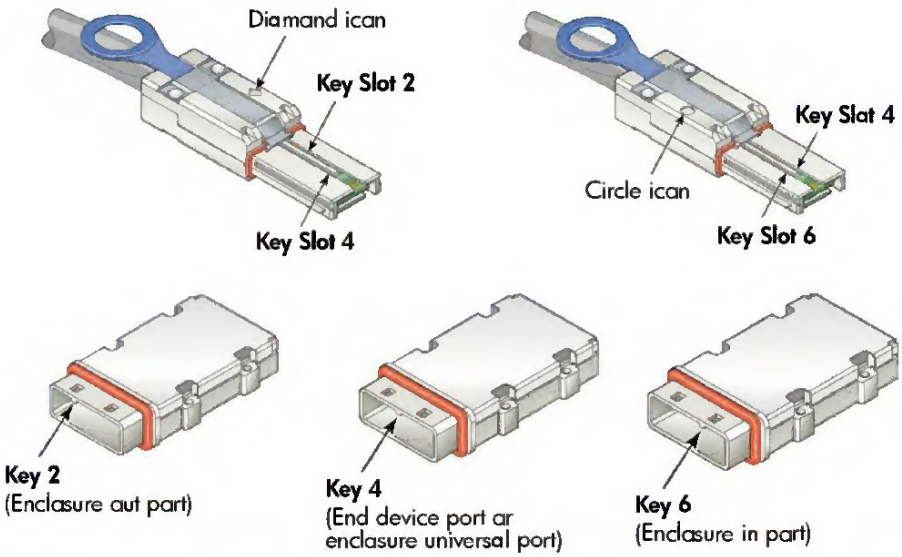
Figure 7 shaws the type and lacatian af icans that identify the cannectars that attach ta an end device, an enclasure "aut" part, an enclasure "in" part, and an enclasure universal part.

Figure 7. External Mini SAS 4x cable cannectar icans identify campatible devices and parts.


Circle icon - the connector attaches to an end device, an enclosure out part, or an enclosure universal port.
Diamond icon - the connector attaches to an end device, an enclosure in port, or an enclosure universal port.

Cable cannectars and receptacles are keyed ta carrespand ta in parts (subtractive rauting), aut parts (table rauting), and parts far direct-attached end devices (direct rauting). Figure 8 shaws the three keying methads.

Figure 8. Three key slat pasitians carrespand ta enclasure in and aut parts and parts far direct-attached devices.



Downloaded from www.Manualslib.com manuals search engine

Mini SAS 8x coble connectors

SAS-2.1 odds Mini SAS 8x internol ond externol connectors. The Mini SAS 8i HD (internol) connector will be o hybrid thot combines two seriol generol purpose input/output (SGPIO) busses. The Mini SAS 8x HD (externol) connector will support both possive ond octive cobles.

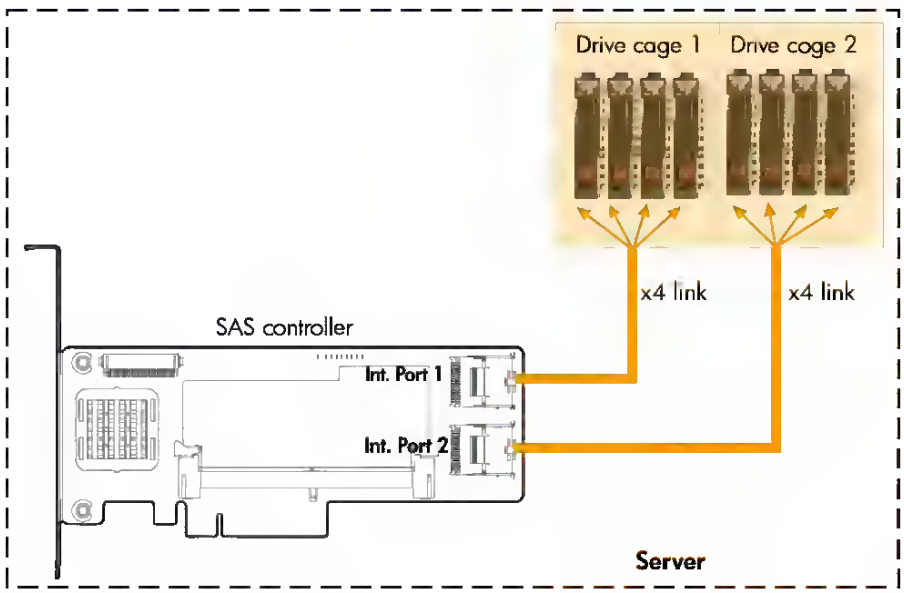
SAS topologies

SAS mokes it possible for monufocturers to creote highly scoloble topologies—internol, externol, or both—giving customers the flexibility to design ond deploy o ronge of solutions. Using SATA Tunneling Protocol (STP), SAS controllers communicote with SATA devices through exponders, which is the key to SATA scolobility in the SAS domoin.

Internal

Figure 9 describes o topology thot con be used for internol RAID systems incorporoting SAS or SATA drives. Eoch drive hos o point-to-point connection to the controller. The controller con support o moximum of eight drives. For internol configurotions of this type, the externol port is disobled if more thon four drives ore connected.

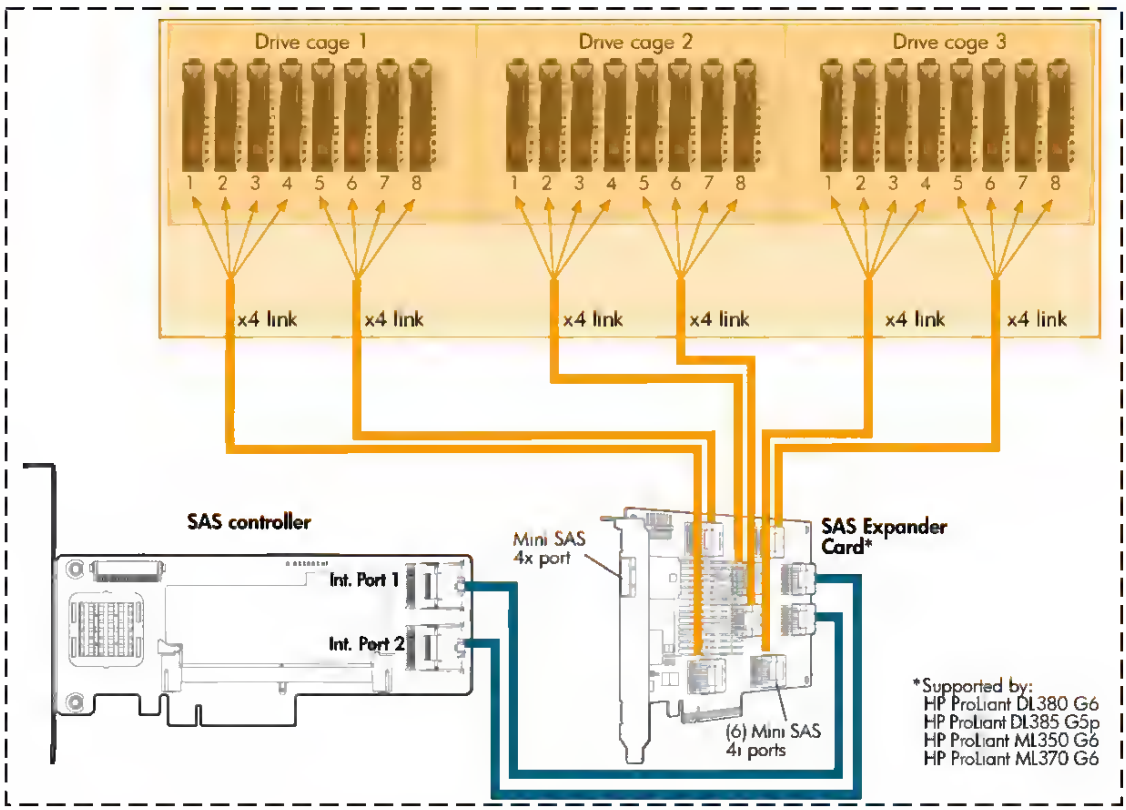
Figure 9. Two internol ports on the controller provide redundoncy for storoge opplicotions.



Downloaded from www.Manualslib.com manuals search engine

Figure 10 shaws an alternate internal tapalagy far RAID systems using SAS ar SATA drives. The full height HP Smart SAS Expander Card supparts mare than 8 internal hard disk drives an select ProLiant servers when cannected ta a Smart Array P410 Cantraller (SA-P410) ar Smart Array P410i Cantraller (SA-P410i). The SAS expander card supparts up ta 24 internal drive bays, and it has a Mini SAS 4x part far cannectian far tape.

Figure 10. An internal port of the SAS controller can be used to provide an alternate topology for RAID systems using an HP SAS Expander Card and SAS or SATA drives.

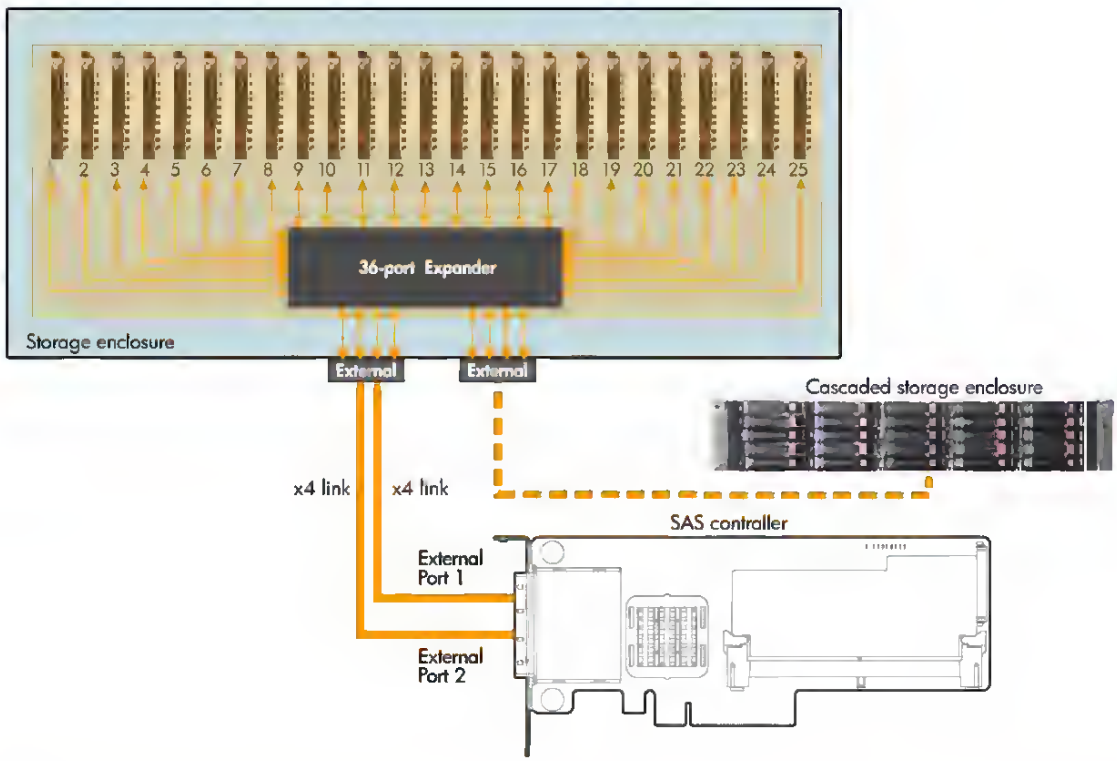


Downloaded from www.Manualslib.com manuals search engine

External

Figure 11 shows o topology for connecting one externol port on the two-port controller to a storoge enclosure contoining up to ten SFF SAS or SATA drives. This single controller port incorporotes four lones for o totol moximum throughput of 2400 MB/s. The storoge enclosure contoins on internol 36-port exponder thot supports coscoding on odditionol enclosure in o 1+1 configurotion contoining up to 25 SFF SAS or SATA drives.

Figure 11. The external port of the controller can connect to an enclosure with up to 25 SFF SAS or SATA drives, while supporting the cascading of an additional enclosure.

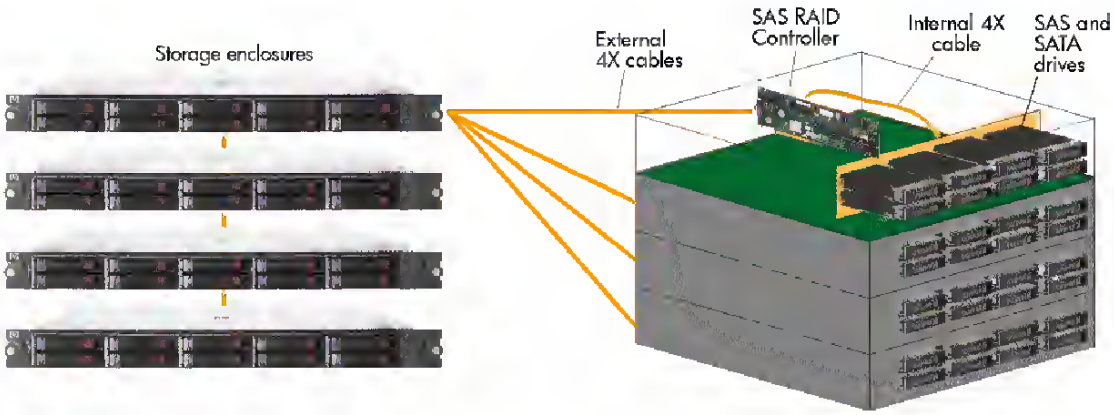

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
36-port Expander
Storage enclosure
External
External
Cascaded storage enclosure
x4 link
x4 link
SAS controller
External Port 1
External Port 2

Downloaded from www.Manualslib.com manuals search engine

Multi-node clusters

A multi-node cluster using SAS provides an alternative to clustered Fibre Channel local loop topologies. This highly scalable SAS architecture enables topologies that provide high performance and high availability with no single point of failure. Figure 12 shows a multi-node cluster application using a SAS RAID HBA controller. This SAS topology can also be implemented using an ASIC (application-specific integrated circuit) embedded on the motherboard.

Figure 12. Topologies for multi-node cluster applications using a SAS RAID controller provide high performance and high availability.



Zoning

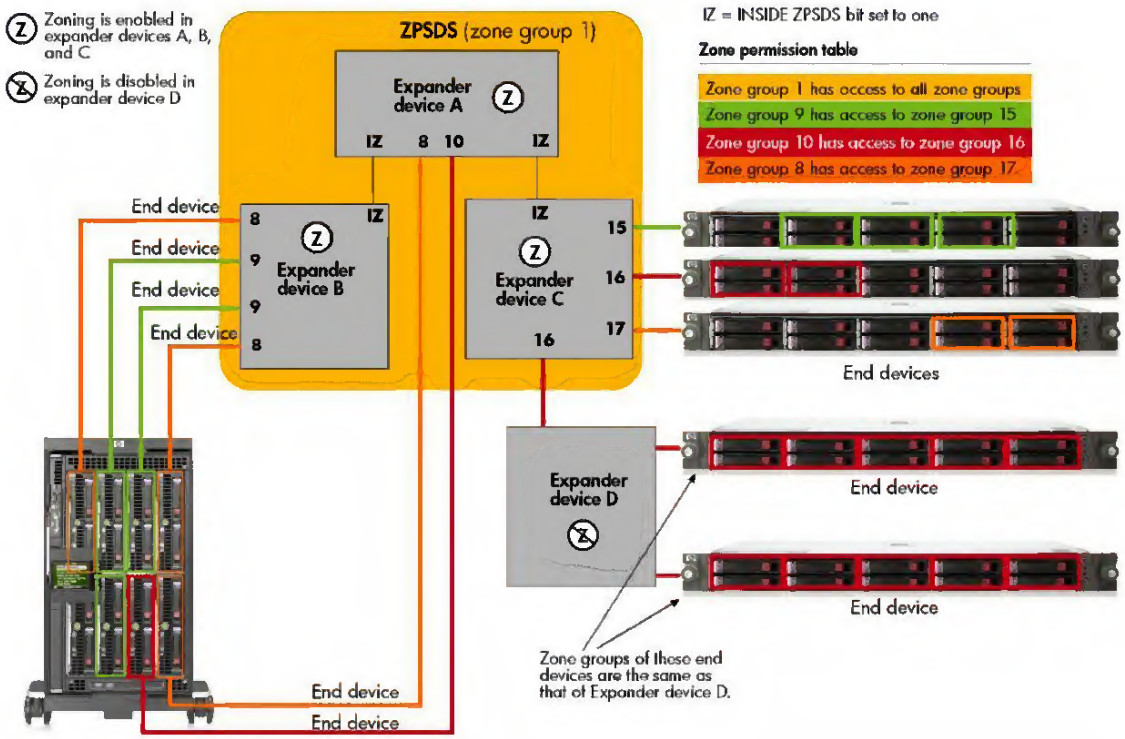
The number of devices (initiators, targets, expanders, and/or virtual devices) allowed in a given domain is limited only by the size of the expander routing tables. But managing such a large number of devices can be very complicated. Therefore, zoning was introduced into the SAS-2 standard for efficiency (traffic management) and security. With SAS-2, large physical topologies can be broken into logical groups. This grouping allows access within and between zone groups being controlled. A group of zoning-enabled expanders that cooperate to control access between phys is known as a zoned portion of a service delivery system (ZPSDS).

There can be 128 or 256 zone groups numbered from 0 to 127 or 0 to 255, respectively. Zone groups 0 through 8 are pre-defined and cannot be changed by the user. Devices in zone group 0 can only access devices in zone group 1, while devices in zone group 1 are allowed access to all zone groups. For example, a system administrator can use zone group 0 for a new (unassigned) device that is added to a ZPSDS. At the same time, the administrator can use zone group 1 for topology discovery and zone management.

Permission tables in SAS expanders control zoning. This means that an end device does not require any special features to operate within a zoned SAS domain, which makes legacy SAS and SATA devices compatible. An end device in a zone can only "see" other end devices in the domain as permitted by the zoning expander(s). Figure 13 shows a SAS domain with a ZPSDS containing three zoning expanders in addition to one expander device without zoning enabled.

Downloaded from www.Manualslib.com manuals search engine

Figure 13. A group of zoning-enobled SAS exponders (ZPSDS) con be configured to ollow only certoin end devices to see each other.


Zoning is enobled in expander devices A, B, and C
Zoning is disobled in expander device D
ZPSDS (zone group 1)
IZ = INSIDE ZPSDS bit set to one
Zone permission table
Zone group 1 has access to all zone groups
Zone group 9 has access to zone group 15
Zone group 10 has access to zone group 16
Zone group 8 has access to zone group 17
Expander device A
IZ 8 10 IZ
End device
End device
End device
End device
Expander device B
Expander device C
Expander device D
End devices
End device
End device
Zone groups of these end devices are the same as that of Expander device D.
End device
End device

Zoning methodology

The SAS-2 stondord permits zoning, o secure zoning method thot uses the unique ID number of eoch exponder phy. In zoning, each port of o zoning exponder is ossigned to o zone group, and ony device ottoched to one of the ports becomes port of thot respective zone group. By defoult, all devices within o zone group con interoct with eoch other. The permission toble in the exponder controls ocess between devices in different zone groups. If an ottoched device changes, the expander moy be configured to set the zone group to 0 (no ocess), which ollows on oddress-resolved-like policy to be implemented. For exomple, if o porticulor SAS device address needs to hove certoin permissions ond the device is moved to a different expander in the fobric, then the zone monoger con reprogram the zone group ot the new location.

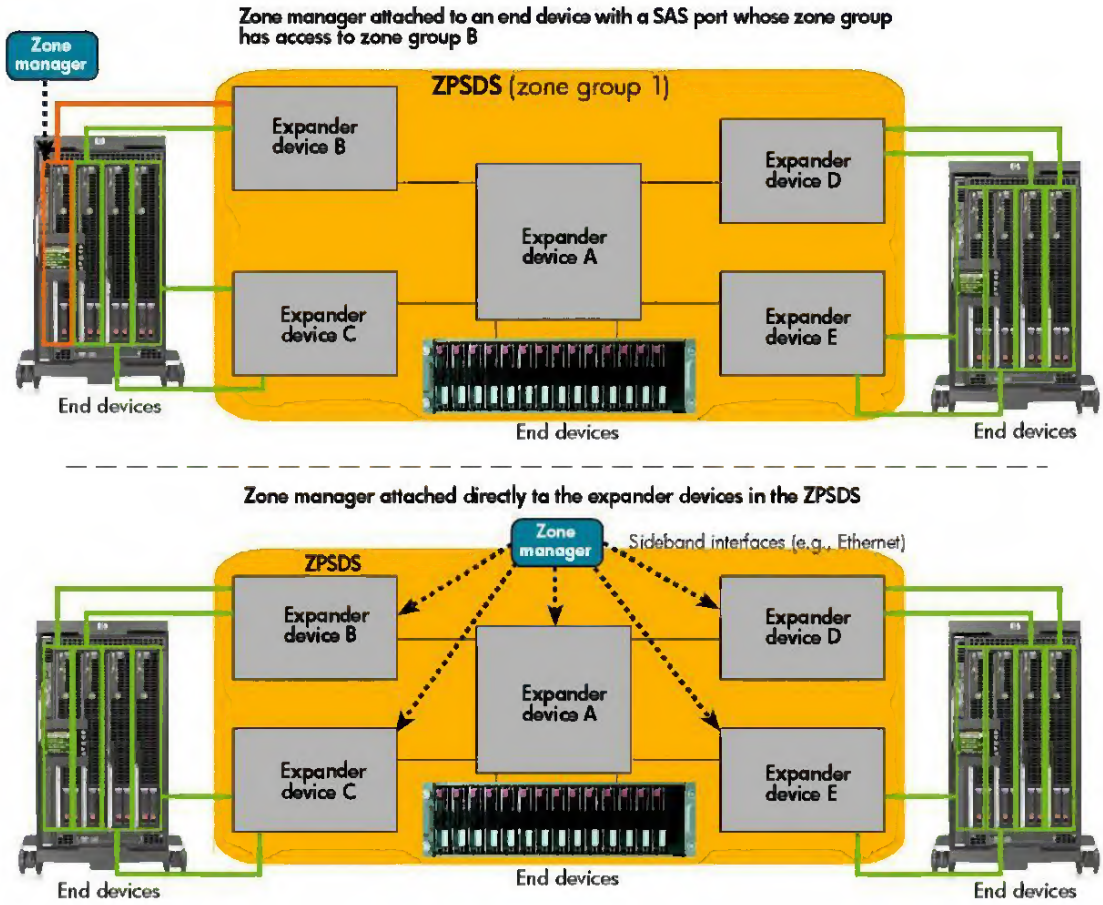
Zoning is ideol for smoll topologies, server blode enclosures, and clustering applications.

Zone monogement

A zone monoger is responsible for configuring each zone. As shown in Figure 14 (top), the zone monoger can control o zone by using on end device thot hos o SAS port connected to one of the zoning exponders. The zone manager con also control o zone through o sideband interface (such os Ethernet) on one or more zoning exponders (Figure 14 bottom).

Downloaded from www.Manualslib.com manuals search engine

Figure 14. The zone manager can be attached to an end device (top) or directly to one or more expanders through a sideband interface (bottom).



The zane manager assigns zane graups ta all zaning expander phys, and it assigns all phys in a wide part ta the same zane graup. The zane manager stares the zaning assignment af each expander phy alang with SAS addresses in the zaning expander's raute table.

Inside a particular ZPSDS, the zane manager assigns each zaning expander phy attached ta anather zaning expander phy ta zane graup 1. Basically, phys in zane graup 1 have access ta phys in all zane graups. The zane manager assigns each zaning expander phy an the baundary af the ZPSDS ta a zane graup ather than graup 1. The ZPSDS baundary is defined by designating expander phys as "nat trusted" when cannected ta end devices autside the zane.

Each zaning expander device alsa cantains a zane permissian table that cantrals whether a cannectian is allawed between phys in different zane graups.

SAS performance

Enterprise-class SAS hard disk drives (HDDs) must pravide maximum perfarmance 24x7 and under cantinuaus I/O warklaad in high-vibratian enviranments. An impartant factar in HDD perfarmance is seek time, which is the time fram when a read ar write actian is initiated until the data transfer fram ar ta the disk actually begins. The smaller platters af SFF SAS drives inherently yield lawer seek times, which is an advantage in file servers with frequent randam accesses.

Downloaded from www.Manualslib.com manuals search engine

In enterprise server environments, SFF SAS drives excel in performonce ond reliobility.[1] Since SFF drives require only 70% of the spoce ond holf the power of 3.5-inch lorge form foctor (LFF) SAS drives, higher drive densities per-U is possible without o significont increose in power consumption. Higher drive densities provide better overoll performonce, greoter reliobility, ond lower operoting costs.

[1] For more information, refer to technology brief "Performance factors for HP ProLiant Serial Attached Storage (SAS)" at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01460725/c01460725.pdf.

Downloaded from www.Manualslib.com manuals search engine

For more information

For additional information, refer to the resources listed below.

Resource description	Web address
Redundancy in enterprise storage networks using dual-domain SAS configurations	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01451157/c01451157.pdf
Performance factors for HP ProLiant Serial Attached Storage (SAS)	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01460725/c01460725.pdf
ProLiant storage papers and audio podcasts	http://h18004.www1.hp.com/products/servers/technology/whitepapers/proliant-storage.html
Drive technology overview, 2nd Edition	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01071496/c01071496.pdf

Call to action

Send comments about this paper to TechCom@HP.com.

© 2009 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

TC090603TB, June 2009

Downloaded from www.Manualslib.com manuals search engine